

A METHODOLOGY FOR EVALUATING THE COSTS AND BENEFITS OF EMPLOYEE ASSISTANCE PROGRAMS

Michael T. French
Gary A. Zarkin
Jeremy W. Bray

Employee assistance programs (EAPs) represent a relatively new fringe benefit for workers; the number of these programs has been steadily increasing in worksites of all sizes. Despite this surge in the growth of EAPs, few studies have estimated their costs or benefits. To guide future economic evaluation studies of EAPs, we have developed a methodology that has four components: a process description to understand the structure, operating environment, and goals of the EAP; a cost analysis to comprehensively identify and estimate the full range of EAP costs; an outcomes analysis to rigorously estimate the effectiveness of the EAP for groups of employees and the overall impact of the EAP on employee performance and workplace productivity; and an economic evaluation to estimate cost-effectiveness ratios, dollar benefits, and net benefits of the EAP. Our methodology is based on standard economic theory, but we present the evaluation strategy in a nontechnical way so that it can be used by employers and other researchers to estimate the costs and benefits of EAPs.

Introduction

An increasing number of worksites have established employee assistance programs (EAPs) for workers with a variety of personal problems. A Bureau of Labor Statistics (BLS) survey found that 6.5% of establishments offered EAP services in 1988 while a follow-up survey of those same worksites found that 11.8% had an EAP in 1990 (Hayghe 1991). A more recent study by the Research Triangle Institute (RTI) found that 33% of all private nonagricultural worksites with 50 or more full-time employees currently offer EAP services. Furthermore, an additional 9% of those worksites in the RTI study not currently offering EAP services plan to start an EAP within the next year (Hartwell et al. 1994). Looking at coverage from the employee perspective, another study found that 45% of a

Michael T. French, Ph.D., is a senior health economist with Research Triangle Institute and an adjunct professor with the Graduate Public Policy Program at Georgetown University. Dr. French has published many papers on the economics of addiction and evaluations of drug abuse interventions. He is currently leading a large evaluation study of EAPs. Gary A. Zarkin, Ph.D., is a senior economist and program director of the Health and Human Resource Economics program at the Research Triangle Institute. Dr. Zarkin is a distinguished labor economist with considerable experience in substance abuse research and program evaluation. His research is published in several leading economics journals. Jeremy W. Bray, M.A. is a research economist at the Research Triangle Institute and a doctoral candidate in economics at the University of North Carolina at Chapel Hill. Mr. Bray has participated in several program evaluation studies and he has particular expertise in econometric analysis and database management. Address questions and reprint requests to Michael T. French, Center for Economics Research, Research Triangle Institute, 1615 M Street, NW Suite 740, Washington, D.C., 20036.

sample of full-time employees had access to an EAP in 1991 (Blum, Martin, and Roman 1992).

As the prevalence of EAPs increases, so does the need for scientifically rigorous studies to evaluate these programs. The primary objective of this article is to provide methodological guidance for future evaluation studies of EAPs. Much of the insight and lessons developed by the authors of this article were obtained through case study analyses of seven EAPs over the last three years. We are in the process of describing the EAP structure, goals, and services; estimating EAP costs and outcomes; and conducting an economic evaluation at each of the worksites. We are challenged in this effort by many of the methodological issues that are outlined in this article. In particular, data collection is always easier in theory than it is in practice; some of the key outcome variables are unavailable at some sites. Also, we must continue collecting data for up to three years at some sites to establish an adequate follow-up period. We hope that the points outlined in this article will help to improve the execution and quality of future EAP evaluations.

Much of our work has been directed by the findings of past EAP evaluations. Some of these studies focus on whether EAPs effectively identify and recruit employees with alcohol abuse, drug abuse, and mental health problems (Levy et al. 1980; Foote and Erfurt 1981; Harris and Fennell 1988; Blum, Martin, and Roman 1992). Other studies have examined the outcomes of a specific type of intervention that complements the core EAP services such as follow-up, constructive confrontation, or various treatment options (Foote and Erfurt 1991; Trice and Beyer 1984; Walsh et al. 1991). However, few studies have examined both the costs and the associated benefits of an EAP itself, independent of any ancillary services.

Separating the EAP effect from the full effect of EAP services coupled with ancillary services poses a formidable challenge to EAP evaluators. Klarreich (1990) describes several confounding factors and logistical difficulties obstructing EAP benefit-cost analyses, such as collecting and valuing employee absenteeism, turnover, medical claims, productivity, and work behaviors, as well as lengthy delays before these improvements are observed. Many researchers discuss the organizational difficulties that often interfere with EAP evaluations, such as the confidentiality of program records, skepticism about the research on the part of administrators, and the constraints faced by EAP staff in implementing the research study protocols (Foote and Erfurt 1991; Trice and Beyer 1984).

Several EAP evaluation studies have been criticized for a poor study design (Kurtz et al. 1984; National Research Council/Institute of Medicine 1994; Bannerman 1992). Experienced researchers in the area of EAP evaluation have recognized that several factors limit the data necessary to implement an experimental design methodology. Although no study has randomly assigned EAP-eligible employees to a control group and an EAP-treated group, some research has involved random assignment to various treatment options *after* contact with the EAP (see Walsh et al 1991; Foote and Erfurt 1991). However, without a control group who obtained no EAP services, inferences about whether outcome changes are attributable to EAP services or to some other factors are difficult to justify. Another factor limiting the analyses of EAPs is the brevity of the observation period. To accurately detect a change in a particular outcome, it is often preferable to have data for several years before and after EAP contact.

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

The time and expense of including a lengthy data collection period precludes this option for many research projects.

Two of the most widely cited evaluations of the independent EAP effect explained above were conducted at McDonnell Douglas Corporation (MDC) (1990) and Control Data Corporation (CDC)(1990). Each of these evaluations has several strengths, but data limitations and methodological problems reduce the statistical significance of the findings. For example, the MDC evaluation is limited by using only ten employees in the comparison group that is matched on demographic characteristics with the EAP treatment group, and by examining only absenteeism and medical claims data. The MDC study is enhanced by a three-year observation period. However, results are insufficiently documented to provide any guidance for future evaluation studies.

The CDC evaluation examines the impact of the EAP on performance reviews, salary increases, absenteeism, and medical claims costs for "troubled" workers. The study includes a comparison group of untreated troubled workers. The evaluators use nine years of data on claims costs, ten years on absenteeism, and eleven years of performance reviews and salary increases. Despite the strengths of a lengthy follow-up period and of an untreated comparison group, the CDC evaluation has limitations. For instance, the study does not report any statistical tests of the main findings or provide the methodology for estimating EAP costs. Also, the study does not use a comparison group of "untroubled" employees to control for any extraneous factors that might be affecting all workers at the worksite.

In addition to the studies noted above, most EAP evaluations tend to be in-house promotional efforts (Cayer and Perry 1988), so they are criticized for potential investigator bias (Krupnick and Pincus 1992). One reason why so few economic evaluations have been conducted is the difficulty in obtaining the necessary cost and outcome data. Much of these data are confidential and the outcome variables (e.g., absenteeism, medical claims) are influenced by many factors other than participation in an EAP.

To guide future evaluation studies, we have identified four components of an evaluation strategy for EAPs:

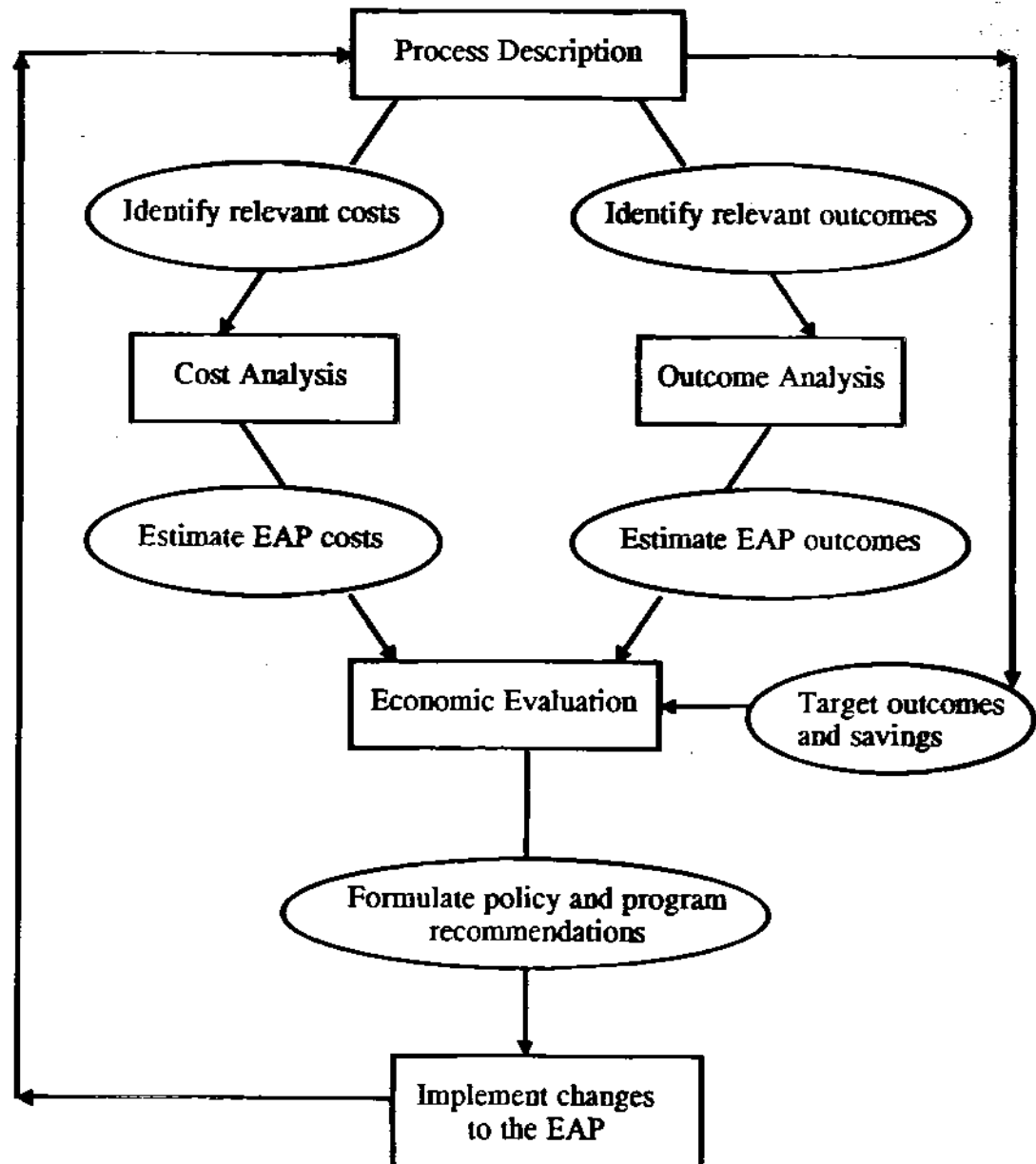
- A process description to understand the structure, operating environment, and goals of the EAP; and to guide the evaluation.
- A cost analysis to comprehensively identify and estimate the full range of EAP costs.
- An outcomes analysis to rigorously estimate the effectiveness of the EAP for groups of employees and the overall impact of the EAP on employee performance and workplace productivity, and
- An economic evaluation to estimate cost-effectiveness ratios, dollar benefits, and net benefits of the EAP.

Although our EAP evaluation strategy contains most of the features of traditional program evaluation strategies found in the social sciences (e.g., Patton 1986), the reader may notice some differences in organization and content. The main reason for these differences is explained by the economic perspective of our program evaluation, which places more emphasis on cost and outcomes analyses.

(For a full description of the economic method of program evaluation see Drummond et al. 1987 or Campen 1986.)

We illustrate the relationship among these four components in figure 1. Each step in the evaluation strategy provides information for and a transition to the next step. The process description identifies important resources and services for the cost analysis and key outcome variables for the outcomes analysis. The cost and outcomes analyses provide estimates for the economic evaluation. The economic evaluation uses findings from the cost and outcomes analyses to estimate cost effectiveness, dollar benefits, and cost savings, which can then be used to develop policy recommendations. The policy and program recommendations can then be used to critique and improve the EAP.

Figure 1
Flow Chart of EAP Evaluation



ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

We organize this article based on our four proposed components of an evaluation strategy for EAPs. Section 2 discusses the process description used to determine program structure, goals, and EAP services delivery. Section 3 describes the cost analysis used to collect financial data and estimate program costs. In section 4 we explain the outcomes analysis, which is used to estimate the effectiveness of the EAP for key outcome variables, and the overall program impact. Section 5 combines the first three components into an economic evaluation framework to estimate cost effectiveness and net benefits of the EAP, and highlight possible policy implications. Section 6 summarizes our methodology and provides some concluding remarks.

Process Description

A process description of the EAP is a critical step to prepare the evaluator for data collection and analysis. Roman and Blum (1988) discuss the structure of EAPs and the importance of a process evaluation. In our EAP evaluation framework, a process study serves the following four purposes:

- Describes the structure of the EAP.
- Describes the environment in which the EAP operates
- Identifies the goals and objectives of the EAP.
- Identifies cost elements and key outcome variables.

The first step, a description of the EAP, should include a history of the EAP, the year the EAP started, any important changes in the operation or structure of the EAP, and the EAP's utilization rates and staffing structure. Given the high degree of variability among EAPs, the process description should document the types and intensity of service provided by the program. Evaluators should also collect information pertaining to outreach activities, counseling services, referral, follow-up, administrative activities, and any other activities that would describe the EAP process. The evaluator should discuss the philosophy of the EAP, the EAP's position in the corporate hierarchy, and whether the EAP is provided internally by company staff or externally by a contractor.

The second step is a description of the operating environment of the EAP, which involves the worksite, the organization, and the community. These environments all play a role in the structure and success of the EAP. The worksite and organizational characteristics include workforce demographics, type and location of facilities, type of industry, corporate structure, and company awareness and cooperation. Community characteristics such as demographic profile, employment patterns, philosophy of local government, and community treatment options all contribute to the structure and success of the EAP. For the process description, the evaluator should collect information about these environments to help explain the operating characteristics and constraints of the EAP.

The third step of the process description identifies the goals of the EAP. Most EAPs provide a wide range of services to employers and employees as described by Roman and Blum's "core technologies" (Roman and Blum 1988). The specific implementation of these core technologies varies from program to program. Therefore, it is necessary to identify the individual goals of the EAP under study.

For example, some EAPs provide an array of counseling and referral services for mental health problems or chemical dependency. Other EAPs strive to reduce the firm's health insurance costs by serving as a gatekeeper to mental health and substance abuse treatment. Still other EAPs make referrals to treatment for employees who test positive for illicit drugs, assist supervisors with the correct way to handle chronically unproductive employees, and organize trauma support groups for employees exposed to workplace violence. Because many, if not most, EAPs have some combination of these and other goals, the evaluator should identify the full range of EAP services before conducting the actual evaluation.

The fourth step in the process description is to identify the important cost and outcome variables for the evaluation. Table 1 presents examples of possible cost and outcome variables for typical EAPs. Cost elements include direct expenditures on personnel and equipment, as well as indirect costs such as shared facilities and equipment and volunteer labor. The cost of an externally provided EAP probably includes only the direct cost of the service contract, while an internal EAP incurs both direct and indirect costs. Outcome variables include absenteeism, health insurance claims, and accident rates. For example, an evaluation of a program designed to reduce unexcused absences and work stoppages would certainly examine absenteeism records and performance reviews for evidence of improved attendance, while an evaluation of a program that serves a managed care role would examine benefits utilization records for evidence of decreased health care costs.

Table 1
Potential Costs and Outcomes of an EAP

Cost Variables	Outcome Variables
Direct Costs	Absenteeism
Personnel	Days absent due to injury or illness
Equipment	Days absent without approval
Supplies	Days tardy
Contractors' fees	Total days absent for any reason
Electricity	Health Care Expenditures
Telephones	Number of outpatient visits
Other utilities	Number of inpatient hospital stays
Staff training	Total charges for outpatient care
Indirect Costs	Total charges for inpatient care
Office space	Salary and Rate of Pay
Shared facilities	Workplace Accidents
Shared equipment	Number of workplace accidents or injuries
Donated time	Days away from work due to workplace accidents or injuries
Volunteers	Total restricted days due to workplace accidents or injuries
	Dollar damage to property and equipment
	Job Performance
	Total score on a performance rating scale
	Promotions for exemplary performance
	Demotions for poor performance
	Number of formal grievances

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

Cost Analysis

The main emphasis of the cost analysis is collecting the necessary data to estimate the total and average costs for the EAP. When collecting cost data, the evaluator should use records of actual expenditures for direct costs, not budgets, to determine the amount spent on a certain cost category. Because budgets are ex-ante projections of future expenditures, they often deviate from actual expenditures, whereas ex-post financial statements reflect past, and therefore, realized costs.

The evaluator should be careful to include all relevant opportunity costs when collecting expenditure data. Opportunity costs represent the value of a resource in its next best use. In most cases the opportunity cost and the dollar expenditure for a resource are identical, but when considering volunteer services or office space used "free of charge," opportunity costs become important. Usually, opportunity costs can be calculated by determining the amount the EAP would pay for the resources if they were not provided free of charge.

Once all the relevant resource-use and expenditure data have been collected, the evaluator can then compute total and average cost estimates for the EAP. Total cost is calculated by simply adding all appropriate cost elements for a given time period. Average costs are then calculated by dividing the total cost estimate by some measure of eligibility or utilization. One set of possible average cost estimates is the cost per eligible program participant (e.g., employees plus dependents) or the cost per eligible employee. To calculate an average cost per eligible person, the evaluator would simply divide the total cost estimate by the number of individuals eligible to use the program. To calculate the average cost per eligible employee, the evaluator would divide total cost by the number of employees eligible to use the program.

Each of the average cost estimates discussed above will describe the costs of the EAP in a slightly different way. For example, since many EAPs cover dependents, as well as employees, the evaluator may consider the effect of the EAP on dependents to be an important aspect of the program. Consequently, the average cost per eligible person would be an appropriate estimate to compare to program outcomes. If, on the other hand, the EAP does not cover dependents or the evaluator considers that effect to be less important, the average cost per eligible employee would be the more appropriate statistic to calculate. The appeal of these estimates is that they provide a per-person cost, which can be compared to other EAPs regardless of size. In addition, the average cost estimates can be compared to other worksite variables such as average insurance premiums, average fringe benefit cost, and average days absent per employee.

Average cost estimates based on eligibility are not always pertinent to the research, however, because they spread the full cost of the EAP over some people who have not used the program. If this is a concern, the evaluator should compute either the average cost per person served or the average cost per employee served. The distinction between these two estimates is the same as between average cost per eligible person and per eligible employee, but the averages for served individuals consider only people who have actually used the EAP. These averages are useful when the evaluator needs to determine the average cost of serving troubled employees as opposed to the average cost of providing EAP services to all employees.

Table 2 uses hypothetical EAP cost and utilization values to illustrate the calculation of average cost per eligible employee and average cost per employee

served. First, we note the total cost of the EAP, which is the sum of all relevant opportunity costs during the test year. Next, we list the number of employees eligible to use the EAP and the number of employees who actually sought counseling from the EAP. Finally, we calculate and display the average cost estimates.

Table 2
Total and Average Costs of A Hypothetical EAP

Category	Amount	Description
Total Cost	\$18,640.00	Sum of all relevant costs of the EAP including indirect costs in the base year.
Eligible Employees	950	Total number of employees eligible to use EAP services in the base year.
Employees Served	79	Total number of employees who actually used EAP services in the base year.
Average Cost per Eligible Employee	\$19.62	Total cost divided by the number of eligible employees.
Average Cost per Employee Served	\$235.95	Total cost divided by the number of employees served.

Although calculating EAP costs may appear to be a relatively straightforward exercise, evaluators should be aware of certain challenges. First, many EAP budget and expenditure reports are subsumed under those of larger departments. Although detailed record keeping may help with this problem, the evaluator must often develop imputation rules to assign expenses to the EAP. Another challenge relates to the collection of financial data from handwritten and hard-copy formats. A detailed record keeping system minimizes some of the potential problems created by handwritten budgets and expense reports, but computerized records often make data collection much easier and more efficient. Finally, the evaluator must be careful and creative when estimating opportunity costs. Imputation rules and assumptions which form the basis of opportunity cost estimates should be clearly stated so that the cost estimates are understandable and conducive to future analysis.

Outcomes Analysis

The outcome study is perhaps the most challenging of the three evaluation components due to the difficulty in measuring the incremental effect of EAP services on outcome variables. To accurately distinguish the effect of EAP services on job performance and other variables, the evaluator must complete at least the following set of tasks:

- Choose a study design and methodology to guide the outcomes analysis.
- Determine the appropriate employee sample(s) to study.
- Establish a time frame for the analysis.
- Collect the necessary data
- Estimate the effectiveness of the EAP for groups of employees.

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

- Estimate the overall impact of the EAP.

The distinction between the last two tasks noted above relates to the scope of the analysis. For example, the effectiveness analysis will estimate changes in a target outcome (e.g., absenteeism) for a certain group of employees who obtained services from the EAP. The cohort of employees in this example can be all individuals who met with the EAP during a given time period or a subset of this group (e.g., male employees who visited the EAP for an alcohol-related problem). Alternatively, the impact analysis will determine the overall success of the EAP *vis-à-vis* the target outcomes and program goals. The impact analysis takes a more global perspective on outcomes and measures changes for the entire population of eligible employees.

To more fully describe each of these tasks, we divide this section into three subsections. First we explain the study design, methodology, employee samples and time frame. Next, we discuss data collection, effectiveness estimation and impact evaluation. Finally, we describe some challenges and limitations associated with the outcomes analysis.

Study Design, Methodology, Samples, and Time Frame

The methodological strengths of experimental study designs are well established. The importance of avoiding selection bias, omitted variable bias, and confounding factors as well as controlling for data limitations in order to ensure statistical validity and generalizability of the research findings is also commonly recognized (see Cook and Campbell 1979; Campen 1986; Patton 1986; Drummond et al. 1987). Another important factor is providing appropriate qualifications concerning the probable occurrence of these conditions during data analysis.

The recognized and acknowledged ideal methodology to estimate the impact of an EAP would be an experimental design that randomly assigned individuals to an intervention group and a control group. With this design, some number of eligible employees would be randomly assigned to an intervention group (e.g., offered EAP services) and an equal number of eligible employees would be randomly assigned to a nonintervention control group (e.g., no access to EAP services). The incremental effect of EAP services on outcome variables of interest could then be estimated by simply subtracting the mean changes between pre-intervention and post-intervention outcomes for the control group from the mean changes between pre-intervention and post-intervention outcomes for the EAP group. By randomly assigning eligible employees to the intervention or control group, the evaluator would ensure that any measurable differences between outcomes for the two groups are the result of the intervention (EAP services in this context) and not the result of confounding factors (e.g., personal characteristics, a downturn in the economy, or a change in the company's health plan).

Despite the conceptual appeal of an experimental design, most studies of this type are very difficult to implement, especially for a social service like EAPs.¹ To even consider an experimental design to estimate the impact of an EAP, the evaluator must first be able to identify a relatively large sample of "troubled" employees who are eligible to use the EAP. This eligible group of employees must then be randomly assigned to the EAP for assessment and possible referral to treatment or to a control group who would not have access to EAP services.

Identifying an eligible group of employees is itself a daunting task, but withholding EAP access from a group of individuals who may desire these services raises a host of ethical issues. For these reasons, an experimental design is rarely used in EAP evaluations.

In developing a more practical outcome estimation strategy, the evaluator should keep in mind those aspects of the experimental design that make it ideal. The one essential aspect is using a control group that allows the evaluator to distinguish the effect of the EAP from the confounding influence of similar services and other factors. Perhaps the most practical and theoretically acceptable alternative design that achieves this goal is a quasi-experiment with nonequivalent comparison groups.

A quasi-experimental design has an intervention group and a comparison group, but instead of randomly assigning eligible employees to these groups before the intervention takes place, the evaluator places individuals in certain groups based on their own EAP and treatment decisions. The intervention group would consist of people who chose to seek help through the EAP on their own. The evaluator should try to assemble comparison groups similar to the intervention group in every aspect except for the participation in an EAP. A major concern with this type of design is the likelihood that the intervention and comparison groups will not come from the same underlying population. In fact, a nonequivalent comparison group design derives its name from this very concern. The evaluator must carefully choose the intervention and comparison groups to ensure that any inherent differences in the groups do not obscure the EAP effect.

Consider the choice of the intervention group. In most cases, this group consists of everyone who received services from the EAP in a given time period. In some cases, the evaluator may choose a subset of these individuals. The most common subset is an intervention group composed of only employees instead of employees and dependents. Some subsets, however, may be based on a specific presenting problem such as alcoholism or drug abuse. The primary goals of the EAP could help to guide the evaluator in choosing an intervention group. For example, an evaluation of a managed-care-based EAP designed to contain mental health and substance abuse claims will probably use a subset of employees who sought help through the EAP for problems related to mental health and substance abuse.

Next, the evaluator chooses a comparison group or groups. These groups will ideally mirror the characteristics of the intervention group absent the participation in an EAP. The broadest choice for such a group is a random sample of all individuals (or all employees) who are eligible to receive EAP services but have not done so. This group would represent the baseline level of outcomes to which the evaluator can compare EAP clients. A second choice for a comparison group would consist of all individuals (or all employees) who sought help for a personal problem(s), but did not use the EAP. This non-EAP-served group allows the evaluator to partially separate the effect of the EAP from the effect of the EAP coupled with ancillary services (e.g., substance abuse treatment). The evaluator can usually form this group by using health insurance claims data to identify individuals (or employees) who filed claims for substance abuse or mental health services, but did not go to the EAP. A non-EAP-served group of individuals who received substance abuse or mental health services may be a reasonable match to the EAP-served group given the high incidence of mental health and substance abuse problems among EAP clients.²

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

After choosing the methodology and forming the study groups, the evaluator must choose a time frame for analysis. An important factor to consider in this decision is how much time is necessary after EAP contact to capture both short-term and long-term effects of the program. For example, a one-year follow-up after EAP contact may be a sufficient amount of time to capture the short-term effects for some outcome variables such as absenteeism or disciplinary reports. The most important EAP effects for some clients, however, may involve more complex variables such as health status improvements and reduced use of medical care. These longer term outcomes may take several years to materialize. Thus, the evaluator should consider a time frame for analysis that will at least be able to capture the short-term effects of the EAP and possibly the long-term effects if these variables are important to the study.

Another factor to consider is how underlying characteristics of the groups may influence outcomes. To adequately determine how the intervention and comparison groups differ and how these differences might affect outcomes, the evaluator must collect information about the individuals prior to the intervention. Thus, the evaluator must select a time frame that allows for sufficient collection of pre- and post-EAP intervention data. We recommend collecting demographic and outcome data for each individual at least three years prior to and three years following the intervention year (i.e., year of EAP contact for the intervention group) to estimate long-term EAP effects. This usually represents the amount of time necessary to identify most changes in target outcomes, examine trends in key variables over time, and assess whether the EAP effects are stable.

Data Collection and Outcomes Estimation

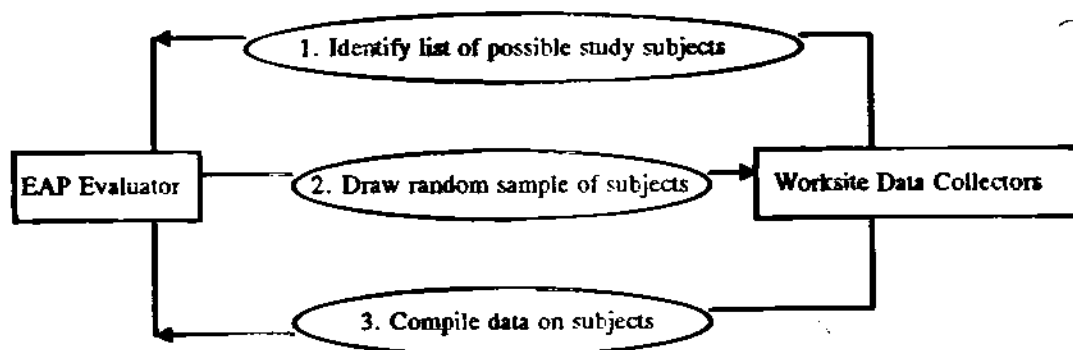
As discussed earlier, key outcome variables are usually identified during the process component of the evaluation. Collection of outcome data at this stage of the analysis should be augmented with information on employee demographics such as age, gender, race and salary. The evaluator must often obtain these data from several different sources and departments before conducting an outcomes analysis. For example, the human resources department may keep absenteeism and demographic data, while an independent insurance company may keep data on medical care utilization and expenditures. Unfortunately, some of the necessary data may be unrecorded or may require labor-intensive record abstraction. For example, few companies keep data on employees for more than a few years, and some companies still keep data in handwritten format instead of electronic files. The evaluator should determine the availability and format of the data before proceeding with the rest of the outcomes analysis. If some data are unavailable or expensive to access, the evaluator may need to modify the analysis to address these deficiencies.

After verifying the availability of necessary data, the evaluator will then draw the employee samples and collect the data for these individuals. Using EAP records, the evaluator can establish the intervention group. Depending on the specific research questions, the evaluator may be able to use personnel records to draw a random sample of employees and health insurance records to select a sample of employees who sought treatment for substance abuse or mental health problems without going through the EAP. After drawing the appropriate employee samples, the evaluator can collect longitudinal data on the outcome variables identified earlier. Because outcome-related data are often confidential and usually kept by separate and distinct departments in the organization, the

evaluator must often rely on staff in those respective departments to compile the data and send it directly to the evaluator. The evaluator could then compile the data and create an analysis file.

When drawing employee samples and collecting outcome data, the evaluator must be especially careful to protect the confidentiality of all individuals involved in the study. We recommend a process similar to the one illustrated in figure 2 to protect employee confidentiality. First, staff in the appropriate departments prepare a roster of possible study subjects and submit these rosters to the evaluator. Second, the evaluator draws a sample of employees from each roster and prepares a master list of all employees selected for the study. The evaluator sends this master list to each of the departments to compile the appropriate data contained in their section. Last, authorized staff in those departments organize those data in a designated format for all subjects and send the data file back to the evaluator to be merged with other files. Although a bit tedious, this process protects employee confidentiality by ensuring that only the evaluator can link an individual to a study group.

Figure 2
Flow Chart of Data Collection Procedure



The last step in the outcomes analysis is to estimate the effectiveness and the impact of the EAP in a statistically rigorous way by controlling for any differences in the intervention and comparison groups or by showing that they are statistically equivalent prior to EAP intervention. To show that the samples are statistically equivalent before EAP contact, the evaluator would compare factors such as age, race, gender, socioeconomic status and education, as well as the outcome variables before EAP intervention. Using statistical tests (e.g., t test or Wilcoxon rank order tests), the evaluator can determine if the groups have statistically equivalent variable means for outcome variables and other factors. If the evaluator cannot reject the null hypothesis of equivalent means, it may be feasible to treat the groups as statistically equivalent. In this case, the evaluator would determine if the EAP was effective by testing whether the difference in an outcome variable between the intervention and comparison groups was statistically different from zero.

In most cases, however, the study groups will not be statistically equivalent. Even if statistical equivalence cannot be rejected based on tests involving

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

observable variables like age, gender and education, individuals differ in many unobservable dimensions that are difficult to detect statistically. When the study groups are not statistically equivalent, the evaluator should develop a framework to describe how observable variables interact to affect employee outcomes. For example, the relationship between an outcome variable and a collection of explanatory variables can be described by equation (1):

$$Y_{t+1} = Y(X_{t+1}, M_{t-1}, M_t, M_{t+1}, EAP_t, R_{t-1}, R_t, R_{t+1}) \quad (1)$$

where

Y = labor market outcome variable (e.g., wage, absenteeism, performance rating).

X = set of demographic variables (e.g., age, ethnicity, gender, education, occupation).

M = set of variables that describe certain classes of personal problems (e.g., depression, hypertension, financial, legal, chemical dependency).

EAP = indicator variable for participation in the EAP.

R = set of variables related to the other types of treatment an individual may receive in addition to, or in place of, EAP services (e.g., outpatient counseling, residential treatment).

t = time period.

The motivation for this theoretical relationship can be explained in the following way. Assume that some individuals receive services from the EAP in time period t . To measure the effect of the EAP on job performance, we want to observe the value of the labor market outcome variable Y in time period $t+1$. The variable Y in period $t+1$ is certainly influenced by a set of demographic variables in time $t+1$, noted by X_{t+1} . Furthermore, Y_{t+1} is affected by the presence of personal problems (e.g., depression or illness) that occur before, during, and after contact with the EAP. This explains why we have the variables M_{t-1} , M_t , and M_{t+1} in the model. Last, EAP clients may receive services from sources other than the EAP that could subsequently affect job performance. These services may include formal substance abuse treatment, psychological counseling, or participation in a self-help group. We recognize the influence of these factors on job performance by including the variables R_{t-1} , R_t , and R_{t+1} in the theoretical model.

Equation (1) posits a basic relationship in which the right-hand-side explanatory variables (X , M , EAP , R) affect the left-hand-side dependent variable (Y). Specified in this way, the evaluator assumes that the explanatory variables are exogenous to the model and are not simultaneously influenced by the dependent variable or other explanatory variables. The evaluator should develop the theoretical model based on the expected relationships among different variables, and not based on the data available.

The evaluator may have already determined what variables are available for analysis, but the theoretical model posits a conceptual framework for how those variables interact. Unfortunately, the theoretical model may not be estimable if some variables are unavailable due to data constraints (e.g., no department kept track of that information) or practical research constraints that limit the resources available to a study (e.g., time and financial expenses might preclude the collection and automation of information that was only kept in handwritten files). Omitted variable bias is the term for the statistical problem that may be created by excluding a variable from the empirical model that should be included based on the theoretical model. Omitted variable bias may be a significant problem if key explanatory variables are missing from the analysis.

Moving to the estimation phase of the study, the evaluator must specify a functional form to define the relationship among the variables noted earlier. The following linear regression model provides a simple and common functional form for EAP outcome estimation:

$$Y_{t+1} = \alpha_0 + X_{t+1}\beta + \alpha_1 M_{t+1} + \alpha_2 M_t + \alpha_3 M_{t-1} + \gamma EAP_t + \alpha_4 R_{t+1} + \alpha_5 R_t + \alpha_6 R_{t-1} + v + \varepsilon_{t+1}, \quad (2)$$

$$Y_{t-1} = \alpha_0 + X_{t-1}\beta + \alpha_1 M_{t-1} + \alpha_2 M_{t-2} + \alpha_3 M_{t-3} + \alpha_4 R_{t-1} + \alpha_5 R_{t-2} + \alpha_6 R_{t-3} + v + \varepsilon_{t-1}, \quad (3)$$

The variables Y , X , M , EAP , and R are the same as defined earlier, the parameters α_i and γ are coefficients to be estimated, and β is a coefficient vector.³ The variable v is an error term that is different for each individual but constant across all time periods and ε is an error term that is constant across all individuals but different in each time period. The t subscripts indicate time periods where period t is the period in which the intervention group sought help from the EAP, $t+1$ is the post-test period, and $t-1$, $t-2$, and $t-3$ are the pre-test periods. The length of time for the pre-test and post-test periods depends on the nature of the study and the outcome variable being examined.

The problem with estimating equation (2) by ordinary least squares (OLS) is that the error term v may be correlated with unobserved client characteristics, which both affect Y (e.g., ambition, motivation, confidence, and discipline) and the decision to go to an EAP. Thus, OLS applied to equation (2) may yield a biased estimate of the effect of going to an EAP. To correct for this potential bias, the evaluator can estimate a "first-difference" model as illustrated in equation (4) (see Moffitt 1991 for a discussion of difference estimators):

$$(Y_{t+1} - Y_{t-1}) = \beta(X_{t+1} - X_{t-1}) + \alpha_1(M_{t+1} - M_{t-1}) + \alpha_2(M_t - M_{t-2}) + \alpha_3(M_{t-1} - M_{t-3}) + \gamma EAP_t + \alpha_4(R_{t+1} - R_{t-1}) + \alpha_5(R_t - R_{t-2}) + \alpha_6(R_{t-1} - R_{t-3}) + (\varepsilon_{t+1} - \varepsilon_{t-1}) \quad (4)$$

By subtracting equation (3) from equation (2), the evaluator eliminates the error component, v , and all other variables that do not change between admission and follow-up. The EAP coefficient γ now represents an unbiased estimate of the effect of EAP services on labor market outcome variables if plausible restrictions on the error terms hold, and if the coefficient vector β and the coefficient estimates α_i are the same for all study groups and across all time periods.

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

Challenges and Limitations

Even with the careful specification of a model, an EAP outcomes study poses at least four challenges to the evaluator. These challenges are

- Time-frame constraints.
- Sample-size limitations.
- Measurement issues.
- Pressure to generate "bottom-line" results.

First, practical considerations such as the cost of collecting data, the availability of historical data, and the necessity of generating a timely report often restrict the length of the evaluation period. Data collection can be extremely expensive, and a longer time frame requires accessing more records, thus increasing the cost of the evaluation. Many worksites maintain data in handwritten files, purge data after a certain period of time, or do not ordinarily compile the requisite data. Lastly, EAPs often evolve at worksites to reflect the personalities of the director and other staff and to meet changing employee needs. Consequently, it is unusual to find a stable EAP model at a particular worksite that has been operating continuously for more than five years unchanged. This poses an especially difficult evaluation challenge because the time frame for pre- and post-EAP analysis is very short.

Data constraints can also limit sample sizes. If a firm has low EAP utilization rates, or if the firm is relatively small, the evaluator is limited in the number of individuals eligible for the intervention and comparison groups. Even if the evaluator considers all individuals who are eligible for the study, low utilization rates in a small firm can leave the evaluator with correspondingly small sample sizes. Small sample sizes for the intervention and comparison groups will reduce the power of the analysis and limit the ability to find statistically significant differences in outcome variables.

Attempting to quantitatively measure the provision of EAP services and other personal care services in the empirical model poses another challenge. The preferable type of variable would measure the amount of services on a continuum from zero to some large amount. The advantage of such a continuous measure is that the evaluator could control for both the provision and intensity of a service. Unfortunately, the evaluator can usually obtain information about the receipt of a service, but data on intensity are poor or nonexistent. Consequently, the evaluator is typically forced to use a dichotomous variable to indicate whether an individual received a service with no additional information about the level or intensity of that service.

Lastly, evaluators are influenced by many factors when determining the objectives and scale of the cost and outcomes evaluation. Quite often, worksite managers and administrators are driven by "bottom-line" results that recommend whether to continue, restructure, or disband an existing program. EAP evaluators may be prematurely subjected to this type of pressure to show that a program is cost effective. Since many EAPs have been operating for only a short time, and a program must be fully implemented before drawing definitive conclusions about cost effectiveness, the evaluation strategy at many worksites may only be capable of developing preliminary recommendations. The point to remember is that the

process description should provide some information and guidance about how to interpret the evaluation results and the recommendations that can be derived from the cost and outcomes analyses.

Economic Evaluation

After completing the cost and outcomes analysis, the evaluator is ready to combine the results from both substudies to conduct an economic evaluation of the EAP. The first step in an economic evaluation is to match the estimated (incremental) costs of the EAP with the corresponding differences in outcomes. Costs are naturally measured in dollars, but outcomes can be measured in other appropriate units, depending on the variable. Dividing incremental cost by the difference for a particular outcome leads to a measure of the cost-per-unit outcome change attributable to the EAP. For example, the evaluator can calculate the cost-per-day-absent avoided for EAP clients by dividing the incremental cost of EAP services by the total reduction in absenteeism for all EAP clients. Comparing this cost-outcome ratio for the EAP intervention with a cost-outcome ratio for an alternative program is called cost-effectiveness analysis. The program with the lowest cost-outcome ratio in this context is judged to be superior for that particular outcome (see, for example, Drummond et al. 1987; French, n.d.; Zarkin et al., n.d.; for a lengthier discussion of economic evaluation techniques). The evaluator can also use these cost-outcome ratios to guide workplace policy by comparing them to other EAPs and other types of programs that may compete for the resources devoted to the EAP.

Matching EAP costs to the corresponding outcomes may seem like a straightforward exercise, but it is critically important. The evaluator should be careful to match outcomes that accrue only to EAP clients with the costs that reflect only the activities of the EAP. The same amount of care should be used when matching costs and outcomes for other types of programs that compete for the same resources. By formally matching the costs and outcomes for each program, the evaluator has the opportunity to examine and compare cost-effectiveness ratios.

Since costs and outcomes are usually measured in different units, cost-effectiveness analysis is somewhat limited for policy purposes. Some of these shortcomings can be avoided by estimating the dollar value of EAP outcomes. When the value of outcomes are measured in dollars and then compared to costs, the exercise is called benefit-cost analysis or net benefit estimation. Benefit-cost analysis and net benefit estimation directly compare the dollar benefits resulting from the EAP to the incremental costs of the EAP. To calculate a benefit-cost ratio, the evaluator would simply divide dollar benefits by costs. For example, the total number of prevented days absent can be assigned a dollar value by determining its wage equivalent. Avoided illness and injury can be valued by determining the medical care expenditures necessary to treat such conditions. The evaluator would then sum the dollar values for all EAP outcomes and divide by the incremental cost of the EAP to calculate the benefit-cost ratio. A ratio greater than one indicates a net savings from the EAP and a ratio less than one indicates that the EAP costs more than it saves. One of the biggest advantages of benefit-cost ratios is that they reveal important information about the EAP without having to compare these ratios to other EAPs or alternative programs.

Whenever feasible, the evaluator should try to compute both cost-outcome and benefit-cost ratios to obtain a complete picture of the EAP's economic impact.

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

Additional policy and program implications regarding the operation of the EAP can also be derived from the results of the economic evaluation. For example, one might ask how the actual savings from the EAP compare to an expected level of savings, or if another program may save more money, is the EAP as effective as desired? Is another program more effective? What are the features of the EAP that make it more or less effective than other programs? Will the EAP model at this worksite be equally effective at other worksites? Answers to these questions will guide the growth of EAPs into cost-effective and beneficial programs for both the sponsoring firm and eligible employees.

Summary

We have presented a conceptual framework and empirical model to guide future evaluations of EAPs. Given the diversity in worksite characteristics and existing EAP models, evaluators should not blindly apply this or any other methodology to all EAPs. Nonetheless, we suggest that all EAP evaluations should try to include at least the following four components:

- Process.
- Cost.
- Outcome.
- Economic evaluation.

When conducting and presenting each of these components, the evaluator should be clear about simplifying assumptions and carefully discuss every possible threat to internal validity. Because each EAP is unique in some way, each evaluation will also be somewhat unique. Nevertheless, the mark of a good EAP evaluation is how closely it adheres to the principles of the conceptual framework, while adapting to fit the characteristics of the program being studied. With the recent growth in EAPs at both small and large worksites, employers and policymakers could gain valuable information from more EAP evaluations that follow the type of methodology proposed in this article.

ACKNOWLEDGMENTS

Financial assistance for this study was provided by two grants from the National Institute on Drug Abuse (NIDA) (contract numbers DA07250 and DA06990). We would like to acknowledge the participation and support of EAP Directors and other company staff at the case study sites. Without their interest and approval, this study would not have been possible. In addition, several colleagues at Research Triangle Institute (RTI) contributed to the study, including Heather Bannerman, Jackie Teague, Ty Hartwell, Frank Potter, and Mike Dennis. Finally, we would like to thank Paul Steele for his advice and suggestions throughout the project. The authors are entirely responsible for the research in this article. It does not represent the position or opinions of NIDA or RTI.

NOTES

1. Walsh et al. (1991) used an experimental design to test the effectiveness of an EAP-based intervention. Walsh and colleagues randomly assigned EAP clients with alcohol-related problems to one of three interventions. They then tested for differences in outcome variables across the three groups. In effect, their design examined the effectiveness of EAP services coupled with a specific type of treatment referral. Our strategy is intended to examine the effectiveness of EAP services only, which makes an experimental design much more difficult to implement because we must withhold EAP services for candidate employees who would be assigned to the control group.
2. A recent study of 439 EAPs found that, on average, about 30% of the caseloads stem from employees' problems with alcohol and other drugs, and 44% involve clients with psychological and emotional problems (Blum et al. 1992; Blum and Roman 1992). It should be noted that these categories are not mutually exclusive, as the EAP clients in this study had, on average, 2.1 problem categories.
3. The variable M is usually difficult to observe and measure for empirical work. The statistical implications of dropping this variable from the empirical model could be significant, but the evaluator may have few alternatives.

REFERENCES

- Bannerman, H. A.
1992 Evaluating the costs and impact of employee assistance programs: Methodological issues and review of selected case studies. Master's thesis, The University of North Carolina at Chapel Hill.
- Blum, T.C., J.K. Martin, and P.M. Roman
1992 A research note on EAP prevalence, components and utilization. *Journal of Employee Assistance Research* 1:209-29.
- Blum, T.C. and P.M. Roman.
1992 A description of clients using employee assistance programs. *Alcohol Health and Research World* 16(2):120-8.
- Campan, J.T.
1986 *Benefit, cost, and beyond: The political economy of benefit-cost analysis*. Cambridge, Mass.: Ballinger Publishing Company.
- Cayer, N.J. and R. Perry
1988 A framework for evaluating employee assistance programs. In *Evaluation of employee assistance programs*, eds. M. Holosko and M. Feit. New York and London: The Hawthorn Press, Inc.
- Control Data Corporation
1990 EAP program evaluation: Employee usage and cost impact 1980-1985.
- Cook, T. D., and D. T. Campbell
1979 *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, Ill.: Rand McNally College Publishing Company.
- Drummond, M.F., G.L. Stoddart, and G.W. Torrance
1987 *Methods for the economic evaluation of health programmes*. Oxford, England: Oxford University Press.

ECONOMIC EVALUATION OF EMPLOYEE ASSISTANCE PROGRAMS

- Foote, A. and J. C. Erfurt
1981 Effectiveness of comprehensive employee assistance programs at reaching alcoholics. *The Journal of Drug Issues* 11(2):217-32.
- Foote, A. and J. C. Erfurt
1991 Effects of EAP follow-up on prevention of relapse among substance abuse clients. *Journal of Studies on Alcohol* 52(3):241-8.
- French, M.T.
n.d. Economic evaluation of drug abuse treatment programs: Methodology and findings. *American Journal of Drug and Alcohol Abuse*. In press.
- Harris, M. M. and M. L. Fennell
1988 Perceptions of an employee assistance program and employees' willingness to participate. *The Journal of Applied Behavioral Science* 24(4):423-38.
- Hartwell, T.D., M.T. French, F.J. Potter, P.D. Steele, G.A. Zarkin, and N.F. Rodman
1994 *Prevalence, cost, and characteristics of employee assistance programs (EAPs) in the U.S.* Research Triangle Park, N.C.: Research Triangle Institute.
- Hayghe, H.
1991 Anti-drug programs in the workplace: Are they here to stay? *Monthly Labor Review*. U.S. Department of Labor, Bureau of Labor Statistics, April 26-28.
- Klarreich, S. H.
1990 The role of cost-benefit analysis in evaluating employee assistance programs. *Compensation & Benefits Management* 6(3):251-5.
- Krupnick, J.L. and H.A. Pincus
1992 The cost-effectiveness of psychotherapy: A plan for research. *American Journal of Psychiatry* 149(10):1295-305.
- Kurtz, N.R., B. Googins, and W.C. Howard
1984 Measuring the success of occupational alcoholism programs. *Journal of Studies on Alcohol* 45:33-45.
- Levy, M. F., W. Reichman, S. Herrington, and D. Young
1980 Alcoholic women in industry: Some empirical data. *International Journal of Group Tensions* 10(1-4):120-9.
- McDonnell Douglas Corporation and Alexander and Alexander Consulting Group
1990 McDonnell Douglas Corporation employee assistance program financial offset study 1985-1989. Bridgeton, Mont. (McDonnell Douglas Corporation) and Westport, Conn. (Alexander and Alexander Consulting Group).
- Moffitt, R.
1991 Program evaluation with nonexperimental data. *Evaluation Review* 15(3):291-314.
- National Research Council/Institute of Medicine
1994 *Under the influence? Drugs and the American work force.* J. Normand, R.O. Lempert, and C.P. O'Brien (eds.). Washington, D.C.: National Academy Press.

FRENCH, ZARKIN, BRAY

Patton, M. Q.

1986 Utilization-focused evaluation. 2d. ed. Beverly Hills: Sage.

Roman, P.M. and T.C. Blum

1988a Formal intervention in employee health: Comparisons of the nature and structure of employee assistance programs and health promotion programs. *Social Science and Medicine* 26:503-14.

Roman, P. M., and T. C. Blum

1988b The core technology of employee assistance programs: A reaffirmation. *The Almacan* August:17-22.

Roman, P.M. and T.C. Blum

1992 Drugs, the Workplace, and Employee-oriented Programming. In *Treating drug problems, Committee for the Substance Abuse Coverage Study*, eds. D.R. Gerstein and H.J. Harwood, Vol. II, 197-244. Institute of Medicine. Washington, D.C.: National Academy Press.

Spicer, J., P. Owen, and D. Levine

1983 *Evaluating employee assistance programs. A sourcebook for the administrator and counselor.* Center City, Minn.: Hazelden Foundation.

Trice, H. M. and J. M. Beyer

1984 Work-related outcomes of the constructive-confrontation strategy in a job-based alcoholism program. *Journal of Studies on Alcohol* 45(5):393-404.

Walsh, D. C., R. W. Hingson, D. M. Merrigan, S. M. Levenson, L. A. Cupples, T. Heeren, G. A. Coffman, C. A. Becker, T. A. Barker, S. K. Hamilton, T. G. McGuire, and C. A. Kelly

1991 A randomized trial of treatment options for alcohol-abusing workers. *New England Journal of Medicine* 325(11):775-82.

Young, D. W., W. R. Reichman, and M. F. Levy

1987 Differential referral of women and men to employee assistance programs: The role of supervisory attitudes. *Journal of Studies on Alcohol* 48(1):22-8.

Zarkin, G.A., M.T. French, D.W. Anderson, and C.J. Bradley

n.d. A conceptual framework for the economic evaluation of substance abuse treatment. *Evaluation and Program Planning.* In press.